Research Report

Dorsal and ventral visual stream contributions to preserved reading ability in patients with central alexia

Oscar M. Aguilar[1,2,5,*], Sheila J. Kerry[3], Jennifer T. Crinion[3], Martina F. Callaghan[2], Zoe VJ. Woodhead [1,2,4] and Alexander P. Leff[1,2,3]

[1]Department of Brain Repair and Rehabilitation; University College London; London, WC1N 3BG; UK.

[2]Wellcome Trust Centre for Neuroimaging; University College London; London, WC1N 3BG; UK.

[3]Institute of Cognitive Neuroscience; University College London; London, WC1N 3AR; UK.

[4]Department of Experimental Psychology; University of Oxford; Oxford, OX1 3UD; UK.

[5]Facultad de Psicologia; Pontificia Universidad Javeriana; Bogota, 110231; Colombia.

* Corresponding author: Oscar M. Aguilar.

E-mail: o.aguilar@ucl.ac.uk

Address correspondence: 12 Queen Square – London, WC1N 3BG. Tel: +44 (0)20 3448 4362

ABSTRACT

We investigated the role of the left temporo-parietal regions in supporting reading abilities of 23 patients with central alexia (CA). For the behavioural data, we employed principal components analysis (PCA), which identified two components: 'reading aloud' and 'reading for meaning'. Voxel-based morphometry of the PCA results showed an association between reading aloud and grey matter density in the left supramarginal gyrus, part of the dorsal visual stream. By contrast, reading for meaning was associated with a large cluster in the left ventral visual stream, from the collateral sulcus to the anterior temporal pole.

Most of the peaks were within the group lesion map, indicating that sparing of these areas results in better preservation of reading ability. However, one white matter (WM) cluster in the medial occipitotemporal lobe was outside the lesioned area. A post-hoc test demonstrated that WM density here was equivalent to controls, suggesting that this was not driven by lesion effects. The two likeliest explanations for this correlation are: 1) that pre-morbid, inter-individual differences in brain structure mitigate the effects of CA; 2) that post-morbid practice-based with reading caused compensatory plasticity. We hope to adjudicate between these explanations with longitudinal therapy data collected in this cohort.

Keywords: Central alexia; Principal component analysis; Supramarginal gyrus; Ventral occipitotemporal cortex; Voxel-based morphometry.

1. INTRODUCTION

Central Alexia (CA) is broadly defined as an acquired reading impairment associated with aphasia, usually caused by middle cerebral artery (MCA) stroke (Leff & Starrfelt, 2014). It usually affects both production (reading aloud) and comprehension (reading for meaning) of written language. Patients with CA will generally be impaired to a greater or lesser extent on different categories of words (usually tested by asking them to read aloud), and will be impaired on tests of other language modalities (such as speaking and writing).

In general, cognitive models of reading propose two complementary routes that allow reading by sound and reading by sight (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Plaut, 2008; Plaut, McClelland, Seidenberg, & Patterson, 1996). Reading by sound (known as the sublexical or direct route) maps between orthography and phonology (O-P) and allows the subject to read novel words, regular words and pseudo-words. Reading by sight (known as the lexical or indirect route) is mediated by semantics (O-S-P), and is used more heavily for words with irregular spellings. In general, reading models propose that both routes are activated simultaneously in reading tasks, but participation of each route varies depending on the type of word (i.e. pseudoword, regular or irregular), feedback received (Plaut et al., 1996) and reading styles (Hoffman, Lambon Ralph, & Woollams, 2015; Patterson & Ralph, 1999; Woollams, Ralph, Plaut, & Patterson, 2007).

Reading deficits in aphasic patients can be classified according to which types of orthographic stimuli patients have problems with. The most accepted classification of acquired reading deficits, referred to by convention as dyslexias, includes three canonical forms: phonological, surface and deep dyslexia. Phonological dyslexia is characterized by poor performance on reading pseudowords and some real words (mostly low-imageability, function and morphologically complex words) with relative preservation of

irregular word reading (Beauvois & Derouesne, 1979). Patients exhibit difficulties in applying grapheme-to-phoneme conversion (GPC) rules and their reading appears to rely more on semantic knowledge of written words (Crisp & Lambon Ralph, 2006). In contrast, patients with surface dyslexia are more impaired when reading irregular words and their reading of regular and pseudowords is reasonably intact (Binder et al., 2016). These patients appear to be relying on their knowledge of the GPC rules, which explains their errors when reading irregular words (e.g. regularisation) (Marshall & Newcombe, 1973; Whitworth, Webster, & Howard, 2014; Woollams et al., 2007). Finally, deep dyslexia refers to patients who have difficulties reading all types of words. They find high frequency, low imageability words (e.g. function words such as [in]) particularly difficult and are more likely to make semantic errors (Jefferies, Sage, & Ralph, 2007; Marshall & Newcombe, 1973; Snowden, Kindell, Thompson, Richardson, & Neary, 2012). These patients show both impaired semantic mediation of word recognition and difficulties using GPC rules during reading. In terms of the reading routes described in cognitive models, phonological and deep dyslexia are said to be caused by damage to the sublexical route meanwhile surface dyslexia is due to damage of the lexical route (Coltheart et al., 2001; Patterson & Ralph, 1999; Whitworth et al., 2014; Woollams, 2014).

Canonical cases of dyslexia have shown that reading processes can dissociate. The extant literature on CA gives the impression that, in clinical practice, one is likely to encounter easily classifiable cases. In our clinical experience, patients do not fall neatly into the classical subdivisions of dyslexia subtypes. An analysis of 64 cases from the PLORAS database (Seghier et al., 2016) found that 78% were 'mixed' (Leff & Starrfelt, 2014). This is in line with other authors who consider phonological and deep dyslexia as a continuum (Crisp & Lambon Ralph, 2006) and surface dyslexia as a more common manifestation of semantic dementia than stroke (Woollams et al., 2007). This has led us to favour a data-driven approach in this study. This approach identifies the componential structure of reading ability over scores from multiple reading tasks. The advantage of this approach rather than relying on individual task scores, is that it is less sensitive to the measurement error (noise) inherent in each task due to task design or the patients' co-commitant non-linguistic impairments.

In the case of anatomical lesion models, extensive research in patients with stroke or dementia has provided a better understanding of brain regions and their role supporting reading skills. Voxel-based morphometry (VBM) and voxel-based lesion symptoms mapping (VLSM) studies have shown that, beyond visual areas, reading skills are mediated by a set of largely left-lateralized brain regions around the territory of the MCA. Ripamonti et al. (2014) found that phonological dyslexia in chronic stroke is related to lesions in the left precentral gyrus, insula, and par opercularis of the IFG whilst surface dyslexia is associated with damage to the left superior, middle and inferior temporal gyri, insula and middle occipital gyrus. Similarly, Binder et al., (2016) found that regularization of irregular words in surface dyslexia is associated to damage in the left posterior middle temporal gyrus. Brambati, Ogar, Neuhaus, Miller, & Gorno-Tempini (2009) studied the anatomical correlates of reading impairments in patients with primary progressive aphasia. They found a correlation between pseudoword reading ability and sparing of the left angular gyrus and posterior middle and superior temporal lobe which are areas associated with the dorsal visual stream (Dehaene et al., 2010); while reading irregular words correlated with sparing of the left temporal pole, anterior middle and superior temporal gyrus, and anterior fusiform gyrus (parts of the ventral visual stream). In the same vein, fMRI studies have shown similar findings linking activation of left parietal regions to phonological processing (Oberhuber et al., 2016) and activation of left temporal regions to semantic processing during reading tasks (Hoffman et al., 2015). In summary, these findings lead to the hypothesis that phonological dyslexia and sublexical O-P reading are reliant on dorsal parts of the MCA territory (inferior parietal lobe, posterior lateral temporal lobe and dorsal inferior frontal cortex), whereas surface dyslexia and lexical O-S-P reading are reliant on ventral MCA areas (ventral temporal lobe and middle-to-anterior lateral temporal lobe).

These previous studies made use of the variability of reading behaviour in a group of patients by correlating behavioural measures on a single test of reading ability with brain regions; however, there are many different ways to assess reading skills and these tests vary in the extent to which they stress O, S, and P processing, as well as motor skills (e.g.: speech output). In this study, we used a novel and complimentary data-driven approach in which we aimed to identify independent cognitive patterns underlying reading

performance in patients with central alexia (CA). By definition CA includes any acquired reading impairments, such as phonological, surface and deep dyslexia. We analysed reading performances of CA patients in different tasks by using Principal Components Analysis (PCA), which is a multivariate technique suitable to reduce large data sets with large numbers of variables while preserving as much as possible of the variance from the original data. PCA identifies patterns in the data by transforming correlated variables into the minimum number of linear components (Field, 2013; Jolliffe, 2002). Thus, it transforms the general structure of the data into a reduced number of components. Recent studies have implemented fruitfully PCA to examine multidimensional aspects of aphasic patients' language abilities (Butler, Lambon Ralph, & Woollams, 2014; Halai, Woollams, & Lambon Ralph, 2017; Lambon Ralph, Snell, Fillingham, Conroy, & Sage, 2010). After identifying the key behavioural components of reading, we aimed to identify which areas of MCA territory showed significant brain-behaviour relationships with the reading components using VBM. Our search area comprised temporal and parietal lobes of the left hemisphere as these encompass the dorsal and ventral streams (Dehaene et al., 2010) thought to support O-P and O-S-P reading respectively. Although the inferior frontal lobe is important for reading, our patients were recruited for a therapy study involving transcranial direct current stimulation (tDCS), which selectively required intact or partially preserved IFG cortex. As these criteria meant that we could not make lesion-deficit inferences about the role of frontal regions in CA it was omitted from our pre-defined search area.

2. MATERIALS AND METHODS

2.1. Participants

23 patients with CA (15 males, mean age 54.4 years, range 25 – 78 years; see Table 1 for demographic details) were recruited from both the PLORAS database (Seghier et al., 2016) and out-patient speech and language therapy services at the National Hospital for Neurology and Neurosurgery, University College London Hospitals. All participants gave written informed consent and the protocol was approved by the London Queen Square Research Ethics Committee at University College London. The CA patients took part in

a longitudinal therapeutic study involving a computer-based word-reading retraining therapy and tDCS to the left inferior frontal gyrus (IFG), the results of which are being prepared for publication. Here we report the patients' baseline (pre-therapy) reading abilities and investigate how these correlated with surviving brain tissue in the left temporo-parietal lobe.

We used the following inclusion criteria: (i) left MCA stroke with preserved or partially preserved left IFG (a requirement of the stimulation therapy); (ii) at least one year post-stroke; (iii) in the aphasic range on either the naming or the spoken picture description subtests of the Comprehensive Aphasia Test (CAT) (Swinburn, Porter, & Howard, 2004) (non-aphasic range for Naming objects subtest= 42-48, inclusion criterion <42; Naming actions range= 8-10, inclusion criterion <8; Spoken picture description range= 33-87, inclusion criterion <33); (iv) impaired on the reading subtest of the CAT (Non-aphasic range= 44–48, inclusion criterion <44); (v) English as their dominant language; and, (v) normal or corrected to normal vision and audition. Participants were excluded if they had: (i) left ACA or PCA stroke; (ii) a history of other neurological or psychiatric condition; (iii) a history of developmental dyslexia; (iv) any type of peripheral alexia; (v) any contraindications for MRI scanning; or, (vi) a severe impairment in speech production based on the word repetition subtest of the CAT (Non-aphasic range= 30–32, exclusion criterion <29). There were no restrictions regarding aphasia or alexia subtype.

For comparison with the broader literature on alexia we have included details of patients' reading profile according to classical alexia subtypes (See Supplementary table 1), although we and others would argue that these criteria are not absolute (Crisp & Lambon Ralph, 2006; Leff & Starrfelt, 2014; Welbourne & Ralph, 2007). Phonological (P), deep (D) or surface (S) dyslexia were categorised using definitions from (Whitworth et al., 2014). Phonological dyslexia was defined according to the presence of a lexicality effect (and possibly an imageability effect) with no regularity effect or semantic errors in word reading. Deep dyslexia was defined according to a lexicality and imageability effect, no regularity effect, but with semantic errors. (Patient 20 showed a weak regularity effect (p=0.024), but his profile was judged to be more similar to deep than surface dyslexia). Surface dyslexia was defined according to a regularity effect, no lexicality effect, but regularization errors (Whitworth et al., 2014).

A lexicality effect was defined as significantly worse accuracy reading non-words than real words using Pearson's chi-square test. Regularity and imageability effects were defined based on binary logistic regressions of the participant's word reading accuracy scores. The predictors used in this regression were word imageability, regularity, length, frequency and N-size. Regularity and/or imageabilty effects were considered significant if the overall model was significant, and the coefficient for the predictive variable (regularity / imageability) was also significant. If the model was not significant, the pattern of word reading errors was used instead."

Patient ID	Age (in years)	Gender	Time post-stroke (in months)	Lesion volume in cm^3	Handedness	Central alexia subtype
1	44	Male	94	240.9	R	D
2	50	Male	82	304.5	R	D
3	64	Male	25	102.7	R	P
4	52	Male	66	122.7	R	P
5	56	Female	93	149.8	R	S
6	55	Female	75	151.2	R	P
7	33	Female	59	181	R	P
8	67	Male	107	11.7	R	D
9	43	Female	55	399.2	R	D
10	61	Male	19	195.6	R	D
11	52	Male	12	31.2	R	P
12	50	Female	14	59.4	R	P
13	54	Male	24	149.3	R	P
14	56	Male	23	45.1	R	P
15	54	Male	39	189.7	R	P
16	73	Male	158	205.2	R	D
17	60	Male	16	102.6	R	D
18	78	Male	22	128.5	L	P
19	50	Female	72	141.3	R	P
20	72	Male	101	243.3	R	D
21	58	Female	41	297.7	R	P
22	42	Male	13	43.7	L	P
23	26	Female	81	161.9	R	D

Table 1. Demographic and clinical information on each patient. R= right; L= Left; P= phonological dyslexia; S= surface dyslexia; D= deep dyslexia.

We collected control data from 23 age - and gender-matched healthy controls (15 males, mean age 54.4, range 23 – 76 years) for tasks without published normative data. An independent samples t-test showed no significant difference in age between the groups (t(44)=.012, *p*=.991). Additionally, a dataset of 29 previously collected healthy control subjects' MRI scans was used to identify patients' brain lesions and for post-hoc comparison of the patients' structural MRI data. This group was also age-matched to the CA patients (18 males, mean age 54.6, range 20 – 72 years; t(50)=-.050, *p*=.960).

2.2. Study design

All patients underwent linguistic assessment over two consecutive sessions, spaced by a two week interval, in which reading and naming tasks were administered.Then, four weeks later (and prior to starting therapy), they had a volumetric MRI brain scan. Healthy controls completed all the reading tasks (see below for list) in one session, but picture naming was not tested because normative data is available for this test (Swinburn et al., 2004).

2.3. Linguistic assessment

Reading abilities were probed using the following tests: (1) word reading; (2) pseudoword reading; (3) written semantic matching; (4) written sentence to picture matching; and, (5) text reading. Additionally, (6) confrontation naming was tested as a measure of aphasia severity.

Word reading: A set of 590 words were selected from the SUBTLEX lexical database (Brysbaert & New, 2009). All words were between three and six letters in length (e.g. accept, bank, bigger) and had high written frequency ($SUBTLEX_{WF}$> 50). Words were presented in a random order and split into six separated blocks, three at each testing session. Words were presented in black, lower case, size 36 Arial font on a grey background using E-prime (Schneider, Eschman, & Zuccolotto, 2012). Participants were instructed to read the words aloud into the voice-key microphone as quickly and accurately as they could. Participants were

given up to four seconds to read the word, responses after this time were scored as incorrect. The experimenter recorded accuracy by button press: 1 for a correct response; 0.5 for a self-correction; and 0 for an incorrect response or failure to respond. Reaction time (RT) was recorded by the voice key. Mean reaction times were calculated excluding: incorrect (or self-corrected) trials where the voice-key did not record the response; and, reaction times more than two standard deviations away from the subject's mean RT".

Pseudoword reading: 20 pseudowords were generated using Wuggy software (Keuleers & Brysbaert, 2010). Stimuli were created from a sample of words used in the word reading task. Pseudowords had three to six letters in length with plausible letter combinations (e.g. tooch), and comparable bigram frequency as the stimuli in the word reading task. The pseudowords were presented in black, lower case, size 36, Arial font on a grey background using E-prime (Schneider et al., 2012). Participants were instructed to read them aloud into a voice-key microphone as quickly and accurately as they could. There was no time restriction for producing a response. Accuracy and reaction time were scored and recorded as described for the word reading test. Similarly, percentage accuracy and mean reaction times were calculated.

Written semantic matching (semantic matching): This task consists of 72 trials presented in E-prime (Schneider et al., 2012) to assess reading for meaning and silent reading speed. In each trial three words were displayed on the screen. Participants read silently a probe word centre aligned at the top and displayed in a white box with magenta contour. Below this probe word there were two words: a semantically-related target (same category or semantically associated) and an unrelated distractor. Probe, target and foil words could be concrete or abstract words (e.g. write, letter, wear). Words were contained in white boxes with blue contours that were left and right aligned. Participants were instructed to decide which word was semantically related to the probe word as quickly and accurately as they could and respond by button press. There was no time restriction for producing a response. A fixation cross was presented for one second between the response and the onset of the next trial. Accuracy and reaction time were recorded by button press (1 point for a correct response; and 0 for an incorrect response). Percentage accuracy and mean

reaction time (for correct trials, excluding trials where reaction time was more than two standard deviations away from the mean) were calculated.

Written sentence to picture matching (sentence reading): This task assessed silent reading for speed and comprehension. It consists of 60 trials, presented in E-prime (Schneider et al., 2012), in which patients silently read a sentence of between five and eight words. Sentences were created using words from the SUBTLEX database (Brysbaert & New, 2009). Moreover, they were matched for word imageability, regularity and frequency. Participants were requested to read each sentence as quickly as they could (e.g. "I like the gold ring"), and to press the space bar once finished. This response was used to determine sentence reading duration. A picture was then displayed on screen and the participant responded verbally whether the picture was congruent with the sentence or not. Half of the pictures were congruent with the sentence (e.g. a happy person looking at a ring), and half were incongruent. The incongruent pictures were based on sentences with similar structure, but different lexical items (e.g. a man enjoying music). Outcome measures were percentage accuracy on the picture decision task and sentence reading speed in words per minute (WPM), excluding trials where speed was more than two standard deviations away from the subject's mean.

Text reading: Level one and two texts from the Neale Analysis of Reading Ability test (Neale, 1997) were administered. Participants were instructed to read the texts aloud as quickly and accurately as possible. Immediately after the participant completed the text, questions were asked to assess reading comprehension. Reading accuracy was recorded for each word. If they could not read a word within four seconds, the experimenter supplied the word and it was scored as incorrect. Self-corrections were scored as correct. Reading accuracy (the percentage of words read aloud), mean reading speed in words per minute, and reading comprehension (questions answered correctly) were calculated.

Naming: As a measure of aphasia severity, patients completed the naming objects (n=24) and actions (n=5) subtests of the CAT (Swinburn et al., 2004). Participants named the object or action in depicted in a line

drawing. Responses were scored as follows; 2- correct response within 5 seconds, 1- correct response after 5 seconds or self-correction error, 0- incorrect response. Total score from both tests were summed (ceiling = 58) and percentage accuracy was calculated.

2.4. Statistical Analysis

Statistical analyses were performed using SPSS (IBM, Released 2013). Firstly, we compared reading performance between patients and healthy controls applying independent sample t–tests. As published norms are available, one sample t–tests were carried out for the naming tasks. Then, a principal component analysis (PCA) of reading scores in the patient group was performed in order to identify common underlying psychological process tapped by these reading tests.

Principal component analysis (PCA)

Principal Component Analysis was used to identify meaningful dimensions underlying performance in the eleven measures of reading (accuracy and speed measures on pseudoword reading, word reading, semantic matching, sentence reading and text reading tests). PCA is a data reduction technique that retains the maximum amount of total variance in the original data by transforming multiple variables into the smallest number of linear and uncorrelated components (Butler et al., 2014; Field, 2013; Jolliffe, 2002).

Before performing PCA, Pearson's correlations were computed to ensure that no pair of reading measures had bivariate correlations stronger than R=0.9 (see Supplementary Material - Table 2), as strongly correlated variables can bias the PCA and should be excluded (Field, 2013). Patients' scores were entered into a PCA with varimax rotation to produce orthogonal factors. This configuration is suitable to study independent components in the data because it produces the minimum number of uncorrelated dimensions conformed by the smallest group of loadings which significantly contribute to the component. This method guarantees that components are unique, therefore results are easily interpretable (Butler et al., 2014; Halai et al., 2017).

Only nine of the eleven reading variables were included in the PCA: pseudoword reading speed was excluded because there were no correct trials to extract reaction times from in 9 out of 23 participants. Also, reading comprehension from the text reading task was no included because its coefficient in the anti-image correlation matrix was .37. Variables with coefficients below .50 may affect components reliability and should be removed (Field, 2013). Factors with an eigenvalue ≥ 1 were used to extract components that account for most of the variance in the data. The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was .645, and the Bartlett's test of sphericity was significant, $X^2(36) = 103.9$, $p<.001$. There is no fixed rule about the adequate sample size for PCA, but in addition to the KMO coefficient, adequacy can be calculated based on subject to variable ratio ≥ 1.2 (Butler et al., 2014; Halai et al., 2017; MacCallum, Widaman, Zhang, & Hong, 1999). In our study we tested 23 patients and we had 9 variables giving a ratio = 2.6, indicative of an adequate sample size.

Five participants had incomplete data (see table 4), representing 4% of the values. Word reading speed could not be calculated for four participants (8, 10, 16 and 20) due to the low number of correctly read words (<20%). Participant 20 was unable to attempt the pseudoword and text reading tasks. Participant 8 was unable to attempt the semantic matching task, and both participants 8 and 17 were unable to attempt the sentence reading task: due to the 2-alternative forced-choice nature of these tests the missing accuracy values were replaced with chance level (50%). All other missing values were addressed using multiple imputation in SPSS by creating ten imputed datasets. Averaged scores of the imputed datasets were used in the PCA.

2.5. Structural MRI data acquisition

All participants were scanned in a 3T whole body MR system (Magnetom TIM Trio, Siemens Healthcare, Erlangen, Germany) equipped with a standard 32 channel head coil for signal reception. Traditionally, Voxel Based Morphometry (VBM) studies use T1 weighted images (T1w). However, in our study we used the magnetization transfer map (MT) from a multi-parameter mapping (MPM) protocol (Weiskopf et al., 2013).

The contrast of this high-resolution MT map offers better grey matter (GM) and white matter (WM) segmentation than standard T1w images, especially in deep subcortical regions (Helms, Draganski, Frackowiak, Ashburner, & Weiskopf, 2009).

We used the same high resolution quantitative MPM protocol described by Callaghan and colleagues (Callaghan et al., 2015). However, we acquired data at 1mm isotropic resolution using a field of view of 256 mm head-foot, 240 mm anterior-posterior, and 176 mm right-left. To accelerate the sequence, partially parallel imaging with an acceleration factor of 2 was used in each of the phase-encoded direction. The GRAPPA algorithm was used with 44 and 40 integrated reference lines in the first and second phase-encoded directions.

MRI Pre-processing and Lesion Mapping

To obtain the MT maps, the MPM images were pre-processed in the Statistical Parametrical Mapping 12 (SPM12; http://www.fil.ion.ucl.ac.uk/spm/), using the Voxel Based Quantification toolbox (http://www.fil.ion.ucl.ac.uk/Research/physics_info/QuantMRI_VBM.html) and Matlab 2014a (The MathWorks, 2014).

The MT maps were spatially normalized into standard MNI space and segmented into GM, WM, cerebrospinal fluid, and an extra lesion tissue-class using the Automatic Lesion Identification (ALI) toolbox optimised for focal brain lesions (Seghier, Ramlackhansingh, Crinion, Leff, & Price, 2008). GM and WM probability maps for the VBM analysis were modulated by the Jacobian determinants of the deformations and smoothed with an isotropic Gaussian kernel of 8 mm full width at half maximum.

Lesions were identified using the ALI toolbox (Seghier et al., 2008) by comparing patients' segmented MT maps to comparable MT maps from a previously collected data set of 29 healthy controls. Control' scans were pre-processed following the same procedure as patient' scans. Next, a binary lesion image was calculated for each patient and then all images were overlapped and thresholded to obtain a lesion overlay

map (LOM) (See supplementary material – Figure 1). The overlay covered perisylvian regions in the left hemisphere corresponding to the anatomical distribution of the MCA. The brain region where the maximum number of patients had damage (n=20) was the WM of the superior longitudinal fasciculus deep to the left supramarginal gyrus (SMG) (x, y, z= -40, -34, 30).

Voxel Based Morphometry (VBM)

VBM was used to identify areas within the left temporal or inferior parietal lobes where GM or WM volume was uniquely associated with the PCA reading components, while controlling for demographic variables and naming ability. GM and WM images were analysed separately. For each analysis, a multiple regression model was created including segmented GM or WM images, PCA reading component scores, plus subjects' age, lesion volume, time post-stroke, and a combined measure of naming scores as nuisance variables. The WFU PickAtlas (Maldjian, Laurienti, Kraft, & Burdette, 2003) was used to create a left temporal and parietal lobe mask. This mask had 2mm 3D dilation to be as inclusive as possible of effects at the edges of the area of interest. The SPM results were interrogated by first selecting a liberal uncorrected voxel (peak) thresholded of $p<0.01$. Then the mask was applied as a small volume correction and only clusters that survived a conservative significance threshold of $p<0.05$ FWE corrected for multiple comparisons were reported. Anatomical regions were labelled using the Harvard-Oxford atlas and the JHU White-Matter tractography atlas (Hua et al., 2008) distributed with FSL (http://www.fmrib.ox.ac.uk/fsl/).

3. RESULTS

3.1. Comparison of Patient and Control Behavioural Data

As expected, the patients' reading performance was significantly worse than the age-matched controls across all nine reading test scores (see Supplementary Material – Table 3); they were also significantly worse than published norms on confrontation naming. This simply confirms that the patients had CA.

3.2. PCA of reading variables

The PCA produced a model with two components with an eigenvalue (e) above 1 (Table 2); the model accounted for 67% of the variance. The first component (e= 4.1) accounted for 46% of the variance and had a high loading on tasks that involved reading words aloud (word reading accuracy, pseudoword reading accuracy and text reading accuracy), hence we labelled this component "reading aloud". The second component (e= 1.9) accounted for 21% of the variance and loaded on tasks involving silent reading for meaning (sentence reading and semantic matching tasks); we labelled this as "reading for meaning". Table 3 summarises patients' scores on the linguistic tasks and PCA components.

	Component Matrix	
	Reading aloud (e=4.1)	**Reading for meaning (e=1.9)**
Text reading accuracy	**.914**	.126
Pseudoword reading accuracy	**.860**	-.232
Word reading accuracy	**.768**	.370
Text reading speed (WPM)	**.733**	.386
Sentence reading accuracy	.271	**.846**
Word reading speed (ms)	.114	**-.729**
Semantic matching speed (ms)	-.257	**-.708**
Sentence reading speed (WPM)	.040	**.654**
Semantic matching accuracy	.455	**.622**

Table 2. Loading of reading tasks on components extracted from the varimax rotated PCA. e= eigenvalue; WPM = words per minute; ms = milliseconds. In bold high loads.

Reaction times for word reading and semantic matching tasks were measured in milliseconds, and so have an inverse association with reading ability (longer reaction times reflect worse reading ability). Conversely, reading speed on sentence and text reading tasks were measured in WPM and so load positively onto the PCA components.

Patient ID	PWR – Acc. (%)	WRT – Acc. (%)	WRT-RT (ms)	SM – Acc. (%)	SM - RT (ms)	SR – Acc. (%)	SR - wpm	TR – Acc. (%)	TR - wpm	Nam-O (Max=48)	Nam – A (Max=10)	Reading aloud	Reading for meaning
1	0	58.35	1377.32	97.22	3708.28	90	84.66	45.33	21.15	35	5	-0.67	0.40
2	0	40.31	1373.07	80.56	4976.5	76.67	73.8	52.56	18.8	28	3	-0.77	-0.28
3	70	96.69	981.88	97.22	1707.83	66.67	35.58	94.66	29.25	38	9	1.39	-0.55
4	0	71.11	791.55	91.67	3431.48	93.33	75.33	34.66	14.45	32	6	-1.00	0.96
5	75	63.82	1956.91	80.56	10854.86	60	28.32	93.58	15	3	0	1.42	-2.65
6	30	91.94	803.76	97.22	3127.52	96.67	61.23	87.17	32.31	46	8	0.66	0.62
7	2.5	90.05	979.44	94.44	5088.54	93.33	37.12	93.58	38.35	46	9	0.51	0.26
8	2.5	12.48	NA	50	NA	50	NA	52	16.63	32	10	-1.22	-1.16
9	20	58.24	1350.99	93.06	1927.03	76.67	50.08	90.67	27.5	42	5	0.44	-0.34
10	0	3.39	NA	51.39	9040.91	41.67	35.74	28	12.39	22	1	-1.43	-1.68
11	75	96.28	872.16	98.61	2072.25	90	91.41	97.44	83.9	41	10	2.00	0.50
12	25	90.59	852.94	95.83	2895.54	90	92.92	94.67	53.84	42	6	0.80	0.78
13	65	91.53	1503.9	97.22	4760.29	96.67	114	98.72	58.28	41	9	1.78	-0.07
14	0	80.37	937.72	97.22	6496.44	86.67	56.42	81.33	20.45	34	8	-0.18	0.32

15	2.5	47.29	1101.92	73.61	4530.96	83.33	32.81	64.1	20.83	6	2	-0.71	-0.26
16	0	19.97	NA	98.61	2161.51	93.33	89.56	11.54	14.66	35	6	-1.22	0.80
17	10	28.14	1256.9	91.67	16336.42	50	NA	30.77	13.29	19	0	-0.86	-1.45
18	7.5	75.42	864.19	98.61	3994.54	93.33	62.21	75.64	48.8	21	4	0.33	0.71
19	5	35.85	757.18	95.83	2251.88	93.33	104.4	67.95	31.01	16	0	-0.82	1.24
20	NA	13.39	NA	95.83	3229.88	76.67	95.61	NA	NA	5	0	-0.04	0.04
21	0	59.49	1138.73	93.06	13351.56	86.67	46.99	70.51	21.19	41	6	-0.48	-0.41
22	27.5	74.92	700.81	95.83	2328.57	90	176.11	86.67	31.89	38	4	-0.01	1.54
23	0	75.51	1249.59	100	2111.24	93.33	94.89	88.46	27.06	41	5	0.08	0.69

Table 3. Patients' scores on the linguistic assessment and PCA components. Abbreviations: PWR= pseudoword reading; WRT= Word reading test; SM= Semantic matching; SR= Sentence reading; TR= Text reading; Nam-O= Naming objects (CAT); Nam-A= Naming actions (CAT); Acc.= Accuracy; RT= Reaction time; ms=milliseconds; wpm= words per minute; Max= maximum score; Comp= component.NA= not applicable.

3.3. Neuroimaging results

VBM analysis

The VBM results (Figure 1) showed positive correlations between the PCA components and tissue volume, controlling for the effects of age, lesion volume, time post-stroke and naming ability. Significant clusters are reported in Table 4.

Component	Cluster size (voxels)	Cluster-level *p* (FWE)	Peak co-ordinate(x, y, z)	Peak location	*Z*
1. Reading aloud (GM) [red region in Figure 1]	533	.05	-44, -44, 36	Supramarginal Gyrus	3.79
2. Reading for meaning (GM) [anterior yellow region in Figure 1]	580	.038	-50, 6, -46	Temporal Pole	3.53
3. Reading for meaning (GM) [posterior yellow region in Figure 1]	777	.011	-52, -52, -10	Inferior Temporal Gyrus	3.33
4. Reading for meaning (WM) [lateral blue region in Figure 1]	1060	.009	-38, -14, -28	Temporal Fusiform Cortex, anterior division	3.90
5. Reading for meaning (WM) [medial blue region in Figure 1]	3316	<.001	-26, -50, -8	WM deep to the collateral sulcus	4.33

Table 4. Anatomical location of brain regions associated to PCA reading components. Regions were determined with the Harvard – Oxford cortical and subcortical structural atlases and JHU White-Matter tractography atlas. GM= grey matter; WM= white matter; FWE= family-wise error correction.

Reading aloud component

Performance on this component was correlated with a cluster of 533 contiguous voxels in the GM of the left SMG (Table 4) extending anteriorly into the parietal operculum. This cluster was wholly contained within the lesioned tissue region; 18/23patients had damage to the peak voxel and 21 had damage to any voxels within this region.

No significant correlations were found between the reading aloud component and WM volume.

Reading for meaning component

Performance on this component was correlated with two GM clusters. The first (777 contiguous voxels) was in the posterior part of the left middle temporal gyrus (MTG) and inferior temporal gyrus (ITG). In this cluster 14 patients had damage to the peak voxel and 15 had damage within the area. The second cluster (580 contiguous voxels) was in the ventrolateral anterior temporal pole. In this cluster no patients had damage to the peak voxel and only 4 had damage within the area. Both of these clusters were located on the boundary of the lesion overlay map.

Voxel values in two WM clusters correlated with the semantic component. Both of these clusters fell mostly outside of the lesioned area. The first cluster was very large (3316 contiguous voxels) and was in the WM extending from left occipital cortex, to left medial temporal cortex. No patients had damage to the peak voxel of the cluster, and as shown in Figure 1 (in blue) the cluster largely fell outside of the lesion overlay area. (However, 20/23 patients had damage to some part of this extensive cluster).The second cluster (1060 voxels) was in WM underlying more anterior portions of the anterior parahippocampal and fusiform gyri. Only two participants had lesioned tissue overlapping the peak voxel of this cluster and 7 had damage within the area.

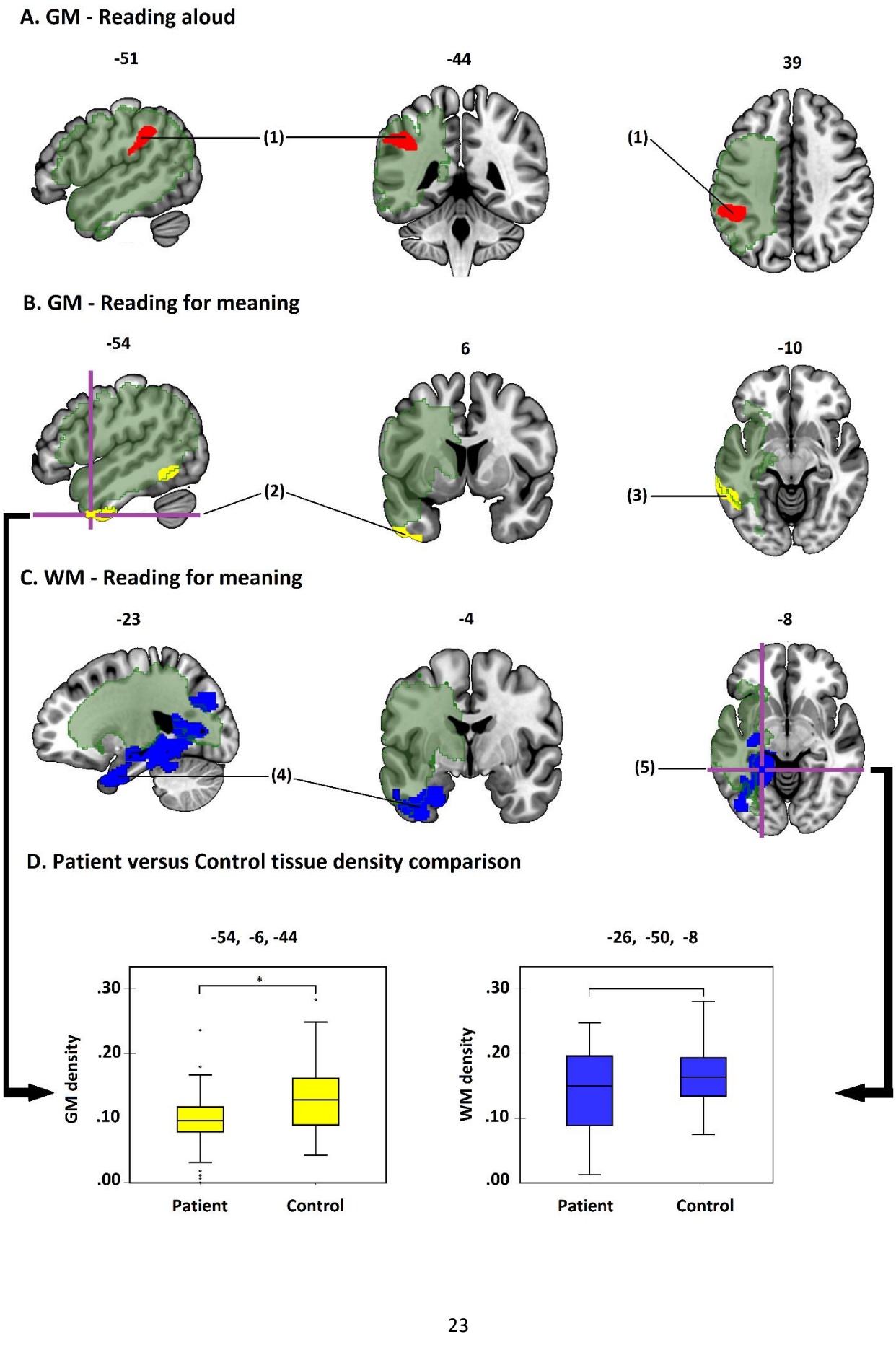
A. GM - Reading aloud
-51
-44
39
(1)
(1)
B. GM - Reading for meaning
-54
6
-10
(2)
(3)
C. WM - Reading for meaning
-23
-4
-8
(4)
(5)
D. Patient versus Control tissue density comparison
-54, -6, -44
-26, -50, -8
GM density
WM density
.30
.20
.10
.00
*
Patient
Control

Figure 1. Voxel-based morphometry results. VBM results show positive correlations between behavioural PCA and tissue volume in grey (GM) and white matter (WM) of the patient group. A: The reading aloud component correlated with a cluster in (1) the left supramarginal gyrus (SMG). B: The reading for meaning component correlated with two GM clusters. One cluster in (2) the left anterior temporal lobe (ATL) and other in (3) the left posterior MTG and ITG. C: The reading for meaning component also correlated with two WM clusters. One cluster covers (4) WM of the middle and posterior left fusiform gyrus. Other cluster encompassing (5) posterior and medial WM deep to the lingual gyrus. The Lesion overlay map (LOM) is shown in green to illustrate clusters within or outside the lesioned areas. Crosshairs in magenta indicates peak co-ordinates at the boarder of/outside the lesioned areas. D: post-hoc analysis of group mean GM and WM densities at these two co-ordinates. Co-ordinates are displayed in X, Y, Z. GM = grey matter; WM = white matter. *= p<.05.

3.4. Post-hoc tests

The region that correlated with the reading aloud component was clearly within the bounds of the group LOM (Figure 1A) as was the first GM cluster in the reading for meaning analysis. Two of the other clusters identified by this analysis were near or outside the borders of the group LOM (Figure 1B and C). We wished to investigate whether these correlations were driven by damage to these regions, either primary or secondary to Wallarian degeneration (in which case the patients as a group should have more damage here compared with controls) or whether the region was not affected either directly or indirectly by stroke (in which case the patients as a group should not differ to controls). To do this we extracted GM and WM density values in the two peak co-ordinates from the patients' images and compared them to the control group's images. A post-hoc unpaired t-test revealed that tissue density in the GM of the anterior temporal lobe was significantly lower in the patient group ($t(50)=-2.3$, $p=.024$). However, we could not reject the null hypothesis for the WM region of the posterior and medial temporal lobe ($t(50)=-1.4$, $p=.167$). See Figure 1D.

4. DISCUSSION

Our aim was to identify brain regions within the left temporal and parietal lobes that correlate with residual reading ability in patients with central alexia. By definition, central alexia is a broad term, which in comparison with alexia subtypes, allows us to study aphasic patients with any reading impairments. Like

other researchers investigating post-stroke aphasia, we chose to utilize a PCA approach to the behavioural data (a form of data reduction) combined with a VBM (mass univariate) analysis of the brain imaging data to identify brain-behaviour correlations (Butler et al., 2014; Halai et al., 2017; Lambon Ralph et al., 2010; Mirman et al., 2015). While mass univariate approaches have been criticized recently as being prone to spatial bias (Mah, Husain, Rees, & Nachev, 2014), the large spatial extent of the regions we identified hopefully mitigate these concerns to some extent. We will concentrate the discussion on our three main findings: (i) PCA analysis of reading behaviour; (ii) VBM results likely related to tissue damage (regions in and at the border of the group lesion overlap map); (iii) VBM results in brain regions where tissue density was largely preserved (regions outside the border of the group lesion overlap map).

The behavioural PCA identified two components, which we labelled as 'reading aloud' and 'reading for meaning'. The first component, reading aloud, explained the largest variance in the data and had high loadings on accuracy of word reading, pseudoword reading, and text reading. This means that our patients presented with a profile characterized by speech production and phonological difficulties (damage to the phonological domain or the O-P connections in the triangle model); this is in keeping with the majority of our patients having a profile consistent with the phonological dyslexia subtype of central alexia (Table 1). In contrast, the 'reading for meaning' component had a high loading on accuracy of written sentence-to-picture matching and written semantic matching tasks. These variables relate to conceptual knowledge, understanding of written words, and text comprehension during silent reading. Furthermore, sentence reading speed in WPM also loaded onto this component showing an association between speed and accuracy in the reading for meaning tasks. It is worth noting that although most patients had phonological impairments (generally referred as phonological dyslexia), the second PCA component indicates that they also showed substantial variation in reading for meaning (damage of the O-S-P route in the triangle model), suggesting some form of semantic deficit. Note that this may not necessarily take the form of classical surface dyslexic reading errors encountered in patients with semantic dementia. Surface errors probably reflect the representational degradation seen in semantic dementia; while in stroke patients executive-semantic impairments are more commonly encountered. (Jefferies, Rogers, Hopper, & Ralph, 2010).

The brain-behaviour VBM analysis within left temporal and parietal lobes revealed that reading aloud correlated independently with one GM region (the SMG, part of the dorsal visual stream), while reading for meaning correlated with two GM and two WM regions (left temporal pole and inferior temporal GM; and ventral occipitotemporal and medial temporal WM; all in the ventral visual stream). The first four regions were all within or at the edge of the patients' LOM while the last region was mostly outside it; the relationship between the identified regions and the LOM affects the inferences that can be drawn. Regions inside the LOM probably support reading behaviour in the undamaged brain, and the degree to which they are spared correlates with residual reading ability. The WM region outside the LOM clearly supports reading but it is probably unaffected, either directly or indirectly by the stroke damage. We will deal with each region in turn.

The reading aloud component identified the GM of the left SMG, region (1). This result supports the idea that the left SMG, which is part of the dorsal visual stream (Dehaene et al., 2010), is crucial in the neural system of reading aloud, linking orthography to phonology; or that it is involved in phonological processing *per se*. Evidence from lesion-behaviour studies of stroke patients with phonological and deep dyslexia (Woollams, 2014) and central alexia (Ripamonti et al., 2014) both identified the left SMG. VBM study in patients with primary progressive aphasia (Brambati et al., 2009) also identified a positive relationship between sparing of the left SMG and phonological reading ability. Studies in controls using fMRI tasks (Graves, Desai, Humphries, Seidenberg, & Binder, 2010; McDermott, Petersen, Watson, & Ojemann, 2003; Oberhuber et al., 2016; Price, 2012), VBM (Carreiras et al., 2009) and transcranial magnetic stimulation (Sliwinska, Khadilkar, Campbell-Ratcliffe, Quevenco, & Devlin, 2012), also strongly support the role of the SMG in phonological processing of written words, and in speech production more generally (Hartwigsen, Golombek, & Obleser, 2015; Mirman et al., 2015).

The remaining four clusters are all associated with the ability to read for meaning and are located in different parts of the left temporal lobe (ventral stream). The first GM cluster (2) is in the ventral anterior temporal

pole (aTL) while the second (3) covers the left posterior MTG and ITG. The anterior temporal pole region has been postulated as a hub that integrates multimodal semantic information for quite some time now (Dilkina, McClelland, & Plaut, 2008; Guo et al., 2013; Hoffman et al., 2015; Lambon Ralph et al., 2010; Patterson, Nestor, & Rogers, 2007; Rice, Lambon Ralph, & Hoffman, 2015). Again, previous VLSM and VBM studies of stroke patients and those with PPA, respectively, have identified both regions as supporting visual semantic processing (Binder et al., 2016; Brambati et al., 2009; Guo et al., 2013; Ripamonti et al., 2014; Wilson et al., 2012). fMRI studies of reading have also shown stronger activation of the left anterior ventral occipitotemporal cortex and posterior part of the middle temporal gyrus in tasks involving the lexico-semantic route (reading irregular words > pseudowords, irregular > regular words, and familiar words > pseudowords) (Price, 2012) and text comprehension (Ferstl, Neumann, Bogler, & von Cramon, 2008).

Two WM regions were identified in our reading for meaning analysis (Figure 1C). The more anterior and lateral cluster (within the LOM, (4)) is large (over 1000 voxels) and covers much of the middle and posterior parts of the left fusiform gyrus, the latter of which consistently demonstrates task-specific activation in many functional imaging studies of single word and pseudoword reading (Price, 2012; Taylor, Rastle, & Davis, 2013; Vigneau, Jobard, Mazoyer, & Tzourio-Mazoyer, 2005; Woodhead, Brownsett, Dhanjal, Beckmann, & Wise, 2011). The posterior fusiform is usually supplied by the posterior cerebral artery, while the middle and more anterior parts are more likely to receive some contributions from the MCA supply. Given that our patients all had MCA territory strokes, it is most likely that the WM of the fusiform features heavily in our LOM because of secondary damage from the initial stroke; that is, Wallerian degeneration that causes deafferentation of regions outside the original stroke area over time. This effect has been observed after large MCA strokes in humans (Gupta et al., 2006).

Perhaps the most interesting finding relates to the final, posterior and medial WM region deep to the lingual gyrus and medial to the fusiform (region 5 in Figure 1C). An association of this region with reading for meaning has been demonstrated by studies on typically developing children (compared with those with developmental dyslexia), where a semantic category judgment task on visually presented words activated

the lingual gyrus (Shaywitz et al., 2002); although meta-analyses of functional imaging studies of reading in normal adults associate this area with lower level visual analysis of written words (Jobard, Crivello, & Tzourio-Mazoyer, 2003). This region is clearly outside the boundary of the LOM and was the only region that had similar tissue density to age-matched control subjects (Figure 1D). The identification of this region cannot therefore be easily explained by any of the mechanisms discussed so far. Two main possibilities arise, both equally compatible with the data presented here: 1) pre-morbid reading ability is related to WM density in this region, so those who have high values here will be less severely affected by their stroke than those who have low values; 2) as the patients are all in the chronic phase (>1 year post-stroke, M=4.7 years) plastic changes in this region (presumably experience-dependent) have occurred since the stroke and support residual reading ability. The first possibility, essentially relating to pre-morbid, inter-individual differences in brain structure that may be driven by genetic or environmental factors, is supported by studies where behaviour in an unselected population correlates with measures of white matter integrity e.g.: fractional anisotropy of posterior white matter correlating with reaction times on a test of visuo-spatial perception (Tuch et al., 2005). The second possibility is supported by evidence from the human expert performance literature where measures of white matter structure correlate with practice-based expertise. In the case of learning to read, this was demonstrated nicely by a study that identified posterior WM tract changes in the splenium of the corpus callosum, after adult illiterates had learnt to read (Carreiras et al., 2009). Hence, the association with reading here may reflect re-modelling of perilesional tracts as a form of post-stroke compensatory plasticity. This hypothesis will be studied further with longitudinal data collected in this cohort.

It is important to note that our findings were biased towards our anatomical area of interest (the left temporal and parietal lobes), chosen because so many other studies had identified these regions as being involved in supporting both reading aloud and reading for meaning. The analysis was not blind to effects in other brain regions but the statistical threshold was higher (FWE corrected for the whole brain volume) and we found no significant regions outside the pre-defined anatomical mask that survived this correction.

5. CONCLUSIONS

PCA of reading abilities in patients with CA has shown a clear dissociation of phonological (reading aloud) and semantic (reading for meaning) dimensions in reading tasks. Additionally, VBM analysis of GM and WM in parietal and temporal regions demonstrated the association of phonological processing with the left SMG in the dorsal stream and semantic processing along the temporal lobe in the ventral stream. This is in agreement with cognitive and anatomical models of reading. Particularly, WM findings highlighted a possible compensatory role of undamaged ventromedial temporal regions in supporting reading ability after stroke, which has not been previous reported.

ACKNOWLEDGMENTS

This work was supported for the Medical Research Council (MR/K022563/1).The Wellcome Trust Centre for Neuroimaging is supported by core funding from the Wellcome Trust 091593/Z/10/Z. The acquisition of the control data was supported by a MRC Clinical Scientist Fellowship (G0701888) awarded to J.T.C.COLCIENCIAS – Administrative department of science, technology and innovation (programme 529, 2011) supported to O.M.A. The authors declare no conflict of interest.

REFERENCES

Beauvois, M. F., & Derouesne, J. (1979). Phonological alexia: three dissociations. *J Neurol Neurosurg Psychiatry, 42*(12), 1115-1124.

Binder, J. R., Pillay, S. B., Humphries, C. J., Gross, W. L., Graves, W. W., & Book, D. S. (2016). Surface errors without semantic impairment in acquired dyslexia: a voxel-based lesion-symptom mapping study. *Brain, 139*(Pt 5), 1517-1526. doi: 10.1093/brain/aww029

Brambati, S. M., Ogar, J., Neuhaus, J., Miller, B. L., & Gorno-Tempini, M. L. (2009). Reading disorders in primary progressive aphasia: a behavioral and neuroimaging study. *Neuropsychologia, 47*(8-9), 1893-1900. doi: 10.1016/j.neuropsychologia.2009.02.033

Brysbaert, M., & New, B. (2009). Moving beyond Kucera and Francis: a critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behav Res Methods, 41*(4), 977-990. doi: 10.3758/brm.41.4.977

Butler, R. A., Lambon Ralph, M. A., & Woollams, A. M. (2014). Capturing multidimensionality in stroke aphasia: mapping principal behavioural components to neural structures. *Brain, 137*(Pt 12), 3248-3266. doi: 10.1093/brain/awu286

Callaghan, M. F., Josephs, O., Herbst, M., Zaitsev, M., Todd, N., & Weiskopf, N. (2015). An evaluation of prospective motion correction (PMC) for high resolution quantitative MRI. *Front Neurosci, 9*, 97. doi: 10.3389/fnins.2015.00097

Carreiras, M., Seghier, M. L., Baquero, S., Estevez, A., Lozano, A., Devlin, J. T., & Price, C. J. (2009). An anatomical signature for literacy. *Nature, 461*(7266), 983-986. doi: 10.1038/nature08461

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: a dual route cascaded model of visual word recognition and reading aloud. *Psychol Rev, 108*(1), 204-256.

Crisp, J., & Lambon Ralph, M. A. (2006). Unlocking the nature of the phonological-deep dyslexia continuum: the keys to reading aloud are in phonology and semantics. *J Cogn Neurosci, 18*(3), 348-362. doi: 10.1162/089892906775990543

Dehaene, S., Pegado, F., Braga, L. W., Ventura, P., Nunes Filho, G., Jobert, A., Cohen, L. (2010). How learning to read changes the cortical networks for vision and language. *Science, 330*(6009), 1359-1364. doi: 10.1126/science.1194140

Dilkina, K., McClelland, J. L., & Plaut, D. C. (2008). A single-system account of semantic and lexical deficits in five semantic dementia patients. *Cogn Neuropsychol, 25*(2), 136-164. doi: 10.1080/02643290701723948

Ferstl, E. C., Neumann, J., Bogler, C., & von Cramon, D. Y. (2008). The extended language network: a meta-analysis of neuroimaging studies on text comprehension. *Hum Brain Mapp, 29*(5), 581-593. doi: 10.1002/hbm.20422

Field, A. (2013). *Discovering statistics using IBM SPSS statistics; and sex and drugs and rock 'n' roll* (4th ed. ed.). Portland: Sage Publications

Graves, W. W., Desai, R., Humphries, C., Seidenberg, M. S., & Binder, J. R. (2010). Neural systems for reading aloud: a multiparametric approach. *Cereb Cortex, 20*(8), 1799-1815. doi: 10.1093/cercor/bhp245

Guo, C. C., Gorno-Tempini, M. L., Gesierich, B., Henry, M., Trujillo, A., Shany-Ur, T., Seeley, W. W. (2013). Anterior temporal lobe degeneration produces widespread network-driven dysfunction. *Brain, 136*(Pt 10), 2979-2991. doi: 10.1093/brain/awt222

Gupta, R. K., Saksena, S., Hasan, K. M., Agarwal, A., Haris, M., Pandey, C. M., & Narayana, P. A. (2006). Focal Wallerian degeneration of the corpus callosum in large middle cerebral artery stroke: serial diffusion tensor imaging. *J Magn Reson Imaging, 24*(3), 549-555. doi: 10.1002/jmri.20677

Halai, A. D., Woollams, A. M., & Lambon Ralph, M. A. (2017). Using principal component analysis to capture individual differences within a unified neuropsychological model of chronic post-stroke aphasia: Revealing the unique neural correlates of speech fluency, phonology and semantics. *Cortex, 86*, 275-289. doi: 10.1016/j.cortex.2016.04.016

Hartwigsen, G., Golombek, T., & Obleser, J. (2015). Repetitive transcranial magnetic stimulation over left angular gyrus modulates the predictability gain in degraded speech comprehension. *Cortex, 68*, 100-110. doi: 10.1016/j.cortex.2014.08.027

Helms, G., Draganski, B., Frackowiak, R., Ashburner, J., & Weiskopf, N. (2009). Improved segmentation of deep brain grey matter structures using magnetization transfer (MT) parameter maps. *Neuroimage, 47*(1), 194-198. doi: 10.1016/j.neuroimage.2009.03.053

Hoffman, P., Lambon Ralph, M. A., & Woollams, A. M. (2015). Triangulation of the neurocomputational architecture underpinning reading aloud. *Proc Natl Acad Sci U S A, 112*(28), E3719-3728. doi: 10.1073/pnas.1502032112

Hua, K., Zhang, J., Wakana, S., Jiang, H., Li, X., Reich, D. S., . . . Mori, S. (2008). Tract probability maps in stereotaxic spaces: analyses of white matter anatomy and tract-specific quantification. *Neuroimage, 39*(1), 336-347. doi: 10.1016/j.neuroimage.2007.07.053

IBM. (Released 2013). IBM SPSS Statistics for Windows (Version Version 22.0). Armonk, NY: IBM Corp.

Jefferies, E., Rogers, T. T., Hopper, S., & Ralph, M. A. L. (2010). "Pre-semantic" cognition revisited: Critical differences between semantic aphasia and semantic dementia. *Neuropsychologia, 48*(1), 248-261. doi: 10.1016/j.neuropsychologia.2009.09.011

Jefferies, E., Sage, K., & Ralph, M. A. (2007). Do deep dyslexia, dysphasia and dysgraphia share a common phonological impairment? *Neuropsychologia, 45*(7), 1553-1570. doi: 10.1016/j.neuropsychologia.2006.12.002

Jobard, G., Crivello, F., & Tzourio-Mazoyer, N. (2003). Evaluation of the dual route theory of reading: a metanalysis of 35 neuroimaging studies. *Neuroimage, 20*(2), 693-712. doi: 10.1016/s1053-8119(03)00343-4

Jolliffe, I. T. (2002). *Principal component analysis* (2nd ed. ed.). New York: New York: Springer.

Keuleers, E., & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. [journal article]. *Behavior Research Methods, 42*(3), 627-633. doi: 10.3758/brm.42.3.627

Lambon Ralph, M. A., Snell, C., Fillingham, J. K., Conroy, P., & Sage, K. (2010). Predicting the outcome of anomia therapy for people with aphasia post CVA: both language and cognitive status are key predictors. *Neuropsychol Rehabil, 20*(2), 289-305. doi: 10.1080/09602010903237875

Leff, A., & Starrfelt, R. (2014). *Alexia : diagnosis, treatment and theory* (Vol. 1). London: London: Springer.

MacCallum, R. C., Widaman, K.F., Zhang S, & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods, 4, 4*, 84-99.

Mah, Y. H., Husain, M., Rees, G., & Nachev, P. (2014). Human brain lesion-deficit inference remapped. *Brain, 137*(Pt 9), 2522-2531. doi: 10.1093/brain/awu164

Maldjian, J. A., Laurienti, P. J., Kraft, R. A., & Burdette, J. H. (2003). An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *Neuroimage, 19*(3), 1233-1239.

Marshall, J. C., & Newcombe, F. (1973). Patterns of Paralexia - Psycholinguistic Approach. *Journal of Psycholinguistic Research, 2*(3), 175-199. doi: Doi 10.1007/Bf01067101

McDermott, K. B., Petersen, S. E., Watson, J. M., & Ojemann, J. G. (2003). A procedure for identifying regions preferentially activated by attention to semantic and phonological relations using functional magnetic resonance imaging. *Neuropsychologia, 41*(3), 293-303.

Mirman, D., Chen, Q., Zhang, Y., Wang, Z., Faseyitan, O. K., Coslett, H. B., & Schwartz, M. F. (2015). Neural organization of spoken language revealed by lesion-symptom mapping. *Nat Commun, 6*, 6762. doi: 10.1038/ncomms7762

Neale, M. D. (1997). *Neale Analysis of Reading Ability - Revised: Manual for Schools.* : Windsor: NFER-Nelson.

Oberhuber, M., Hope, T. M., Seghier, M. L., Parker Jones, O., Prejawa, S., Green, D. W., & Price, C. J. (2016). Four Functionally Distinct Regions in the Left Supramarginal Gyrus Support Word Processing. *Cereb Cortex*. doi: 10.1093/cercor/bhw251

Patterson, K., Nestor, P. J., & Rogers, T. T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nat Rev Neurosci, 8*(12), 976-987. doi: 10.1038/nrn2277

Patterson, K., & Ralph, M. A. L. (1999). Selective disorders of reading? *Current Opinion in Neurobiology, 9*(2), 235-239. doi: Doi 10.1016/S0959-4388(99)80033-6

Plaut, D. C. (2008). Connectionist Approaches to Reading *The Science of Reading: A Handbook* (pp. 24-38): Blackwell Publishing Ltd.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review, 103*(1), 56-115. doi: Doi 10.1037/0033-295x.103.1.56

Price, C. J. (2012). A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. *Neuroimage, 62*(2), 816-847. doi: 10.1016/j.neuroimage.2012.04.062

Rice, G. E., Lambon Ralph, M. A., & Hoffman, P. (2015). The Roles of Left Versus Right Anterior Temporal Lobes in Conceptual Knowledge: An ALE Meta-analysis of 97 Functional Neuroimaging Studies. *Cereb Cortex, 25*(11), 4374-4391. doi: 10.1093/cercor/bhv024

Ripamonti, E., Aggujaro, S., Molteni, F., Zonca, G., Frustaci, M., & Luzzatti, C. (2014). The anatomical foundations of acquired reading disorders: a neuropsychological verification of the dual-route model of reading. *Brain Lang, 134*, 44-67. doi: 10.1016/j.bandl.2014.04.001

Schneider, W., Eschman, A., & Zuccolotto, A. (2012). E-Prime User's Guide. . Pittsburgh: Psychology Software Tools, Inc.

Seghier, M. L., Patel, E., Prejawa, S., Ramsden, S., Selmer, A., Lim, L., Price, C. J. (2016). The PLORAS Database: A data repository for Predicting Language Outcome and Recovery After Stroke. *Neuroimage, 124*(Pt B), 1208-1212. doi: 10.1016/j.neuroimage.2015.03.083

Seghier, M. L., Ramlackhansingh, A., Crinion, J., Leff, A. P., & Price, C. J. (2008). Lesion identification using unified segmentation-normalisation models and fuzzy clustering. *Neuroimage, 41*(4), 1253-1266. doi: 10.1016/j.neuroimage.2008.03.028

Shaywitz, B. A., Shaywitz, S. E., Pugh, K. R., Mencl, W. E., Fulbright, R. K., Skudlarski, P., . . . Gore, J. C. (2002). Disruption of posterior brain systems for reading in children with developmental dyslexia. *Biol Psychiatry, 52*(2), 101-110.

Sliwinska, M. W., Khadilkar, M., Campbell-Ratcliffe, J., Quevenco, F., & Devlin, J. T. (2012). Early and sustained supramarginal gyrus contributions to phonological processing. *Front Psychol, 3*, 161. doi: 10.3389/fpsyg.2012.00161

Snowden, J. S., Kindell, J., Thompson, J. C., Richardson, A. M., & Neary, D. (2012). Progressive aphasia presenting with deep dyslexia and dysgraphia. *Cortex, 48*(9), 1234-1239. doi: 10.1016/j.cortex.2012.02.010

Swinburn, K., Porter, G., & Howard, D. (2004). *Comprehensive aphasia test: CAT*. Hove: Hove : Psychology Press.

Taylor, J. S., Rastle, K., & Davis, M. H. (2013). Can cognitive models explain brain activation during word and pseudoword reading? A meta-analysis of 36 neuroimaging studies. *Psychol Bull, 139*(4), 766-791. doi: 10.1037/a0030266

The MathWorks, I. (2014). MATLAB and Statistics Toolbox Release 2014a. Natick, Massachusetts, United States.

Tuch, D. S., Salat, D. H., Wisco, J. J., Zaleta, A. K., Hevelone, N. D., & Rosas, H. D. (2005). Choice reaction time performance correlates with diffusion anisotropy in white matter pathways supporting visuospatial attention. *Proc Natl Acad Sci U S A, 102*(34), 12212-12217. doi: 10.1073/pnas.0407259102

Vigneau, M., Jobard, G., Mazoyer, B., & Tzourio-Mazoyer, N. (2005). Word and non-word reading: what role for the Visual Word Form Area? *Neuroimage, 27*(3), 694-705. doi: 10.1016/j.neuroimage.2005.04.038

Weiskopf, N., Suckling, J., Williams, G., Correia, M. M., Inkster, B., Tait, R., Lutti, A. (2013). Quantitative multi-parameter mapping of R1, PD(*), MT, and R2(*) at 3T: a multi-center validation. *Front Neurosci, 7*, 95. doi: 10.3389/fnins.2013.00095

Welbourne, S. R., & Ralph, M. A. L. (2007). Using parallel distributed processing models to simulate phonological dyslexia: The key role of plasticity-related recovery. *Journal of Cognitive Neuroscience, 19*(7), 1125-1139. doi: DOI 10.1162/jocn.2007.19.7.1125

Whitworth, A., Webster, J., & Howard, D. (2014). *A Cognitive Neuropsychological Approach to Assessment and Intervention in Aphasia: A Clinician's Guide* (Second ed.). Hove, East Sussex: Psychology Press.

Wilson, M. A., Joubert, S., Ferre, P., Belleville, S., Ansaldo, A. I., Joanette, Y., Brambati, S. M. (2012). The role of the left anterior temporal lobe in exception word reading: reconciling patient and neuroimaging findings. *Neuroimage, 60*(4), 2000-2007. doi: 10.1016/j.neuroimage.2012.02.009

Woodhead, Z. V., Brownsett, S. L., Dhanjal, N. S., Beckmann, C., & Wise, R. J. (2011). The visual word form system in context. *J Neurosci, 31*(1), 193-199. doi: 10.1523/jneurosci.2705-10.2011

Woollams, A. M. (2014). Connectionist neuropsychology: uncovering ultimate causes of acquired dyslexia. *Philos Trans R Soc Lond B Biol Sci, 369*(1634), 20120398. doi: 10.1098/rstb.2012.0398

Woollams, A. M., Ralph, M. A. L., Plaut, D. C., & Patterson, K. (2007). SD-squared: On the association between semantic dementia and surface dyslexia. *Psychological Review, 114*(2), 316-339. doi: 10.1037/0033-295x.114.2.316